RECEIVED

SEQUENCE LISTING

APR 1 7 2002

```
<110> Von Schaewen, Antje
          <120> Plant GntI Sequences and the Use Thereof for the Production ECH CENTER 1600/2900
                 of Plants Having Reduced or Lacking N-acetyl Glucosaminyl
                 Transferase I(GnTI) Activity
           <130> 032266-003
           <140> US 09/591,466
           <141> 2000-06-09
           <150> EP 98/08001
            <151> 1998-09-12
 APR 1 5 2002 2 160> 14
** (210> 1
           <211> 1669
           <212> DNA
           <213> Solanum tuberosum
            <220>
            <221> misc feature
            <222> (659)...(667)
           <223> function: Asn codon in this context is a potential
                  glycosylation site;
                  product: N-glycosylation consensus sequence;
                  phenotype: N-glycans modulate protein properties;
                  standard_name: N-glycosylatoin site;
                  label: pot-CHO;
                  note: GnTI-coding sequences from animals do not
                  contain this feature.
            <221> CDS
            <222> (53)...(1393)
            <223> codon start: 53;
                  function: initiates complex N-glycans on secretory
                  glycoproteins;
                  EC number: 2.4.1.101;
                  product: beta-1,2-N-acetylglucosaminyltransferase I;
                  evidence: EXPERIMENTAL;
                  gene: cgl;
                  standard name: gntI;
                  label: ORF;
                  note: first gntI sequence from potato
                   (unpublished).
             <221> 5'UTR
             <222> (15)...(52)
             <221> 3'UTR
```

<222> (1394)...(1655)

<221> <222> <223>	(80) fund prod star prod note	ction duct: ndare	n: me : hyd d_nar ; dent:	embra droph ne: n	nembi	ane	anch	or (of a	type	e II	Golg	,			
<221> misc_feature <222> (1)(14) <223> function: used for cloning the cDNA library in Lambda ZAPII; product: EcoRI/NotI-cDNA adapter; number: 1.																
<221> <222> <223>	(16 pro	5561.	(1 Ec	_669)	'NotI	-cDN	A ad	lapte	er;							
<400> gaatt	> 1 cgc	gg co	cgcct	tgaga	a aa	cctc	:gaa	ttca	aattt	cg (cattt	ggca	ig aç	g atg Met 1	aga Arg	58
ggg (aac Asn	aag Lys 5	ttt Phe	tgc 1 Cys 1	ttt : Phe :	gat t Asp I	ta d Leu <i>l</i> 10	cgg Arg	tac Tyr	ctt (Leu]	ctc (Leu \	gtc (Val 1	gtg (Val 1	gct g Ala <i>F</i>	gct Ala	106
ctc Leu	gcc Ala 20	ttc Phe	atc Ile	tac Tyr	ata Ile	cag a Gln I 25	atg Met	cgg Arg	ctt Leu	ttc Phe	gcg Ala '	aca Thr	cag Gln	tca (Ser (gaa Glu	154
tat Tyr 35	gta Val	gac Asp	cgc Arg	ctt Leu	gct Ala 40	gct Ala	gca Ala	att Ile	gaa Glu	gca Ala 45	gaa Glu	aat Asn	cat His	tgt (Cys '	aca Thr 50	202
	cag Gln	acc Thr	aga Arg	ttg Leu 55	ctt Leu	att Ile	gac Asp	aag Lys	att Ile 60	agc Ser	cag Gln	cag Gln	caa Gln	gga Gly 65	aga Arg	250
gta Val	gta Val	gct Ala	ctt Leu 70	Glu	gaa Glu	caa Gln	atg Met	aag Lys 75	1110	cag Gln	gac Asp	cag Gln	gag Glu 80	tgc Cys	cgg Arg	298
caa Gln	tta Leu	agg Arg 85	Ala	ctt Leu	gtt Val	cag Gln	gat Asp 90	ДСи	gaa Glu	agt Ser	aag Lys	ggc Gly 95	ata Ile	aaa Lys	aag Lys	346
tta Leu	ato Ile	gga Gly		gtg Val	cag Gln	atg Met 105	FIO	gtg Val	g gca . Ala	gct Ala	gta Val 110	gtt Val	gtt Val	atg Met	gct Ala	394
tgo Cys	agt s Sei		act g Thi	gac r Asp	tac Ty:	_ ren	gag Glu	g ago	g act g Thi	att 11e	1	tcc Ser	ato	tta Leu	aaa Lys 130	442

tac caa aca Tyr Gln Thr	tct gtt g Ser Val A 135	ca tca aa la Ser Ly	S TAT E.	ct ctt t ro Leu E 40	ctc ata tc Phe Ile Se	c cag gat r Gln Asp 145	490
gga tca aat Gly Ser Asn	cct gat g Pro Asp V 150	ta aga aa al Arg Ly	ng ctt g ys Leu A 155	ct ttg a la Leu S	agc tat gg Ser Tyr Gl 16	1	538
acg tat atg Thr Tyr Met 165	cag cac t Gln His L	eu Asp T	at gaa c yr Glu P 70	ct gtg o	cat act ga His Thr Gl 175	aa aga cca Lu Arg Pro	586
ggg gaa ctg Gly Glu Leu 180	gtt gca t Val Ala T	ac tac ac yr Tyr L 185	ag att g ys Ile A	na my	cat tac aa His Tyr Ly 190	ag tgg gca ys Trp Ala	634
ttg gat cag Leu Asp Gln 195	Leu Phe B	cac aag c His Lys H 200	at aat t is Asn E	tt agc Phe Ser 205	cgt gtt a Arg Val I	tc ata cta le Ile Leu 210	682
gaa gat gat Glu Asp Asp	atg gaa a Met Glu 2	att gct g Ile Ala A	rra wah	ttt ttt Phe Phe 220	gac tat t Asp Tyr P	tt gag gct he Glu Ala 225	730
gga gct act Gly Ala Thr	ctt ctt Leu Leu	gac aga g Asp Arg <i>P</i>	gac aag Asp Lys 235	tcg att Ser Ile	atg gct a Met Ala I	tt tct tct le Ser Ser 240	778
tgg aat gad Trp Asn Asp 24	Asn Gly	GIN AIG	cag ttc Gln Phe 250	gtc caa Val Gln	gat cct q Asp Pro P 255	gat gct ctt Asp Ala Leu	826
		ttt cct of Phe Pro 1	ggt ctt Gly Leu	gga tgg Gly Trp	atg ctt 1 Met Leu 3 270	tca aaa tca Ser Lys Ser	874
	c gaa cta r Glu Leu	tct cca Ser Pro 280	aag tgg Lys Trp	cca aag Pro Lys 285	, 1120 - 1 -	tgg gat gac Trp Asp Asp 290	922
	g ctg aaa g Leu Lys 295	gaa aat Glu Asn	cac aga His Arg	ggt cga Gly Arg 300	a caa ttt g Gln Phe	att cgc cca Ile Arg Pro 305	970
gaa gtt to Glu Val Cy	gc aga acg vs Arg Thr 310	tac aat Tyr Asn	ttt ggt Phe Gly 315	OIU III.	t ggt tct s Gly Ser	agt ttg ggg Ser Leu Gly 320	1018
Gln Phe Pl	it aag cag ne Lys Glm 25	tat ctt Tyr Leu	gag cca Glu Pro 330	att aa lle Ly	g cta aat s Leu Asn 335	gat gtc cag Asp Val Gln	1066
		atg gac Met Asp 345	cta agt Leu Ser	tac ct Tyr Le	t ttg gag u Leu Glu 350	gac aac tat Asp Asn Tyr	1114

gtg a Val L 355	aa c ys H	cac i	ttt Phe	ggc Gly	gac Asp 360	ttg Leu	gtt Val	aaa Lys	a aa s Ly	/5 A	ct la 65	aag Lys	cc Pr	c at	tc le	cac His	gg Gl 37	4	1162
gct g Ala A	at d	gct Ala	gtt Val	ttg Leu 375	aaa Lys	gca Ala	ttt Phe	aac Asr	c at n Il 38	re A	at .sp	ggt Gly	ga As	t g p V	u <u>-</u>	cgt Arg 385	at Il	t .e	1210
cag t Gln T	ac Tyr	aga Arg	gac Asp 390	caa Gln	cta Leu	gac Asp	ttt Phe	gaa Glu 39	u A	at a sp I	itc [le	gct Ala	cg Ar	9 -	ag ln 00	ttt Phe	gg G]	ly Ly	1258
att t Ile H	?he	gaa Glu 405	gaa Glu	tgg Trp	aag Lys	gat Asp	ggt Gly 410	va.	a co l Pi	ca o ro P	egg Arg	gca Ala	gc Al 41		at 'yr	aaa Lys	ge G.	ŗλ ââ	1306
ata (gta Val 420	gtt Val	ttc Phe	cgg Arg	ttt Phe	caa Gln 425	Thi	tc Se	t a r A	ga d .rg <i>l</i>	cgt Arg	gtg Val 430		cc c ne I	tt Leu	gtt Val	t t	cc er	1354
cct Pro 1	gat Asp	tct Ser	ctt Leu	cga Arg	caa Glr 440	r reg	gga Gly	a gt y Va	t g il G	Lu.	gat Asp 445	act Thi	t ta	ag (cga	agat	at	g	1403
	agta	atg .	acta	ccac	jta q	gctac	atg	ca a	acat	-cct	aat	ati	t ca.	agt:	t.t.	aca	aac	ataa	1463 1523 1583 1643
<210 <211 <212)> 2 L> 4 2> P	ctt 46 RT	tggc	tube	ege (gaati	.c			. y c c		uuv	Aaa	yya.		900		,	1669
<pre>tgag <210 <211 <212 <213 <400 Met</pre>	(caac) > 2 > 4 2> P 3> S 0> 2 Arg	ett 46 RT olan Gly	tggc ium t	cggco tube n Ly	rosu s Ph	gaati m e Cy	s Ph	e A	sp :	Leu 10	Arc	_ј Ту	r I	jeu Jeu	Lev	ı Va	11 1	Val	1669
<pre>tgag <210 <211 <212 <213 <400 Met</pre>	(caac) > 2 > 4 2> P 3> S 0> 2 Arg	ett 46 RT olan Gly	ium t 7 Asi 1 Ala	tube: n Ly 5 a Ph	rosu s Ph	gaati m	s Ph	e A	sp ln	Leu 10	Arc	_ј Ту	r I	jeu Jeu	Leu	ı Va	11 1	Val	1669
<pre>tgag <210 <211 <213 <400 Met 1 Ala</pre>)> 2 l> 4 2> P 3> S 0> 2 Arg	46 RT olan Gly Let	ium t 7 Asi	tube: n Ly 5	rosu s Ph e Il	m e Cy e Ty	s Ph r Il u Al	e A .e G 2	sp :	Leu 10 Met	Arg	J Ty	r I eu E	eu Phe	Lev Ala	ı Va 15	ıl '	Val Gln	1669
<pre>tgag <210 <211 <212 <213 <400 Met 1 Ala Ser</pre>	(Caac ()> 2 ()> 4 (2> P (3> S ()> 2 ()> Arg () Ala	46 RT olan Gly Lev	num to Asi 20 Asi 20 Asi	tube: n Ly 5 a Ph	rosu s Ph e Il p Ar	m e Cy e Ty g Le	s Ph r Il u Al u Le	e A .e G .a A	sp ln 5	Leu 10 Met Ala	Arc Arc	J Ty g Le e Gl s Il	yr I u I	Jeu Phe Ala	Lev Ala 30 Glv	ı Va 15 a Th	il i	Val Gln His	1669
<pre>tgag <210 <211 <212 <400 Met 1 Ala Ser Cys</pre>	ycaa)> 2 >> 4 2> P 3> S O> 2 Arg Ala Glu	46 RT olan Gly Lev 35 Se: Se:	num to Asin Ala 20 cr Va	tube: n Ly 5 a Ph 1 As	rosu s Ph e Il p Ar r Ar	m e Cy e Ty g Le 55	s Phr Ilu Alu 40	e A e G 2 a A)	sp ln 5 la	Leu 10 Met Ala Asp	Arc Arc Ile Ly	Ty Lee Gl s Il 60 s Hi	yr I Lu I Le S	Jeu Phe Ala 45	Let Ala 30 Glu	Va 15 Th Ass Th	il i	Val Gln His Gln	1669
<pre> tgag <210 <211 <212 <213 <400 Met 1 Ala Ser Cys Gly</pre>	ycaac)> 2 1> 4 2> P 3> S Arg Ala Glu Thr 50 Arg	46 RT olan Gly Let 35 Se: Va.	num d Asi 20 r Va r Gl	tube n Ly 5 a Ph 1 As n Th	rosurs Phe Il PAr Ar A	m e Cy e Ty g Le	s Ph r Il u Al u Le	e A e G 2 .a A) eu I	sp: ln: 5 lla lle	Leu 10 Met Ala Asp Met	Arc	J Ty Lee Gl S Il 60 S Hi	r I I I I I I I I I I I I I I I I I I I	Leu Phe Ala 45 Ser	Let Ala 30 Gli Gli Ass	Value Assumed to the control of the	il 's	Val Gln His Gln Glu 80	1669
<pre> tgag <210 <211 <212 <400 Met 1 Ala Ser Cys Gly 65 Cys</pre>	(Caac)	46 RT olan Gly Let 35 Se: Va.	ium to Asia 20 r Var Gl Van Le	tube: n Ly 5 a Ph 1 As n Th 1 Al 85 Le Gl	rosurs Phe Il PAr Ar A	m e Cy e Ty g Le 55	s Ph r Illu Al 40 40 Le 0	e A e G 22.a A)) eu I lu G al G	sp : ln : 5	Leu 10 Met Ala Asp Met	Arc Ile Ly Ly To Le	J Ty J Le Gl S Ill 60 S Hi u G.	r I Feu F	Leu Phe Ala 45 Ser Gln	Let Ala 30 Gli Ass	1 Va 15 1 Tr 1 As 1 As 1 Pp G. 2 Sp G.	il Voir (Val Gln His Gln Glu 80 Ile	1669
<pre>tgag <210 <211 <212 <213 <400 Met 1 Ala Ser Cys Gly 65 Cys Lys Met</pre>		Gly Let Type 35 See y Value Gles Le	num to Asin Ala 20 r Var Gl Van Le	tube: n Ly 5 a Ph l As n Th l Al 85 c Gl	rosurs Phe Il PAR Ar	m e Cy e Ty g Le g Le g Le j 55 u La La sp Va	s Ph r Ill 40 u Leeu Vi uu G.	e A e G 2 a A lu G al G lu G	sp ln 5 lla lle lle lle lle lle lle lle lle lle	Leu 10 Met Ala Asp Met Asp 90 Pro	Arc Ile Ly 75 Le Va	J Ty J Lee Gl 60 S Hi U G: Ll A:	r I I I I I I I I I I I I I I I I I I I	Leu Phe Ala 45 Ser Gln Ser Ala Ile	Leu Ala 30 Glu Ass Ly Va 11 Ly	1 Va 15 1 Th 1 Ass Th 1 Ass G. 9 9 1 V	il voir of the state of the sta	Val Gln His Gln Glu 80 Ile Val	1669
<pre>tgag <210 <211 <212 <213 <400 Met 1 Ala Ser Cys Gly 65 Cys Lys Met Leu </pre>		46 RT olan Gly Let 35 Se: y Val g Gl: s Le a Cy 11 s Ty	tggc num t Asi 1 Al. 20 r Va r Gl Va n Le 1 10 r Se 5 Gl	tube: n Ly 5 a Ph 1 As n Th 1 Al 2 Ar 85 c Gl 00 ar An 1 Al	rosurs Phe Il PAR Ar	m e Cy e Ty g Le	s Phr Ill 40 Au Le u G: u G: The Triple Trip	e A e G 2a A lu G al G lu G yr I yr I lu S	sp ln 5 lla lle lle lle lle lle lle lle lle lle	Leu 10 Met Ala Asp Met Asp 90 Pro Glu	Arc Arc Ile Ly 75 Le Va Ar	J Ty J Lee Gl 60 S Hi U G: Ll A: TY P	r I I I I I I I I I I I I I I I I I I I	Leu Phe Ala 45 Ser Gln Ser Ala Ile 125 Leu	Lev Ala 30 Glu Ass Ly Va 11 Ly	1 Va 15 1 Th 1 Ass G. 9 9 1 V	il vont	Val Gln His Gln Glu 80 Ile Val Ile Ser	1669
tgag <210 <211 <212 <213 <400 Met 1 Ala Ser Cys Gly 65 Cys Lys Met Let	ycaac 3 2 2 2 2 3 3 5 5 6 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6 7 6	Gly Let 35 Se: Vai Gly Gl: Tyn 37 Tyn	tggc num t Asi 1 Al. 20 r Va r Gl Va n Le 1 10 r Se 7 Gl	tube: n Ly 5 a Ph 1 As n Th 1 Al 20 Eu Ar 50 cer Ar 1 Al 21 cer Ar 21	rosurs Phe Il PAR Ar	m e Cy e Ty g Le 55 eu Gl la Le sp Va er V	s Phr Ill 40 All 40 All GEU V All All All All All All All All All Al	e A e G 2 a A lu G al G f yr I yr I 20 lu S	sp ln 5 lla lle lle lln Met llo5 Leu Ser Arg	Leu 10 Met Ala Asp Met Asp 90 Pro Glu Lys	Arcon Arcon Lys Lys Lee Van Arcon Lys Lee 15	J Ty J Lee Gl S Hi G T T P T P 1 Leu A	rr I u I le S is G hr ro 40	Leu Phe Ala 45 Ser Gln Ser Ala Ile 125 Leu Leu	Lev Ala 30 Glu Ass Ly Va 11 Ly Ph	1 Va 15 1 Th 1 As n GJ 2 S G 9 9 1 V 0 S S S	il vont	Val Gln His Glu 80 Ile Val Ile Ser Gly 160	1669

```
175
                                    170
                165
Arg Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
                                                   190
                                185
            180
Trp Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile
                                                 205
                            200
        195
Ile Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe
                        215
Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
                                         235
                    230
Ser Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp
                                                         255
                                     250
                245
Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
                                                     270
                                 265
            260
Lys Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
                                                 285
                             280
        275
Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile
                                             300
                         295
Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
                                         315
                     310
Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp
                                     330
                 325
Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
                                                     350
                                 345
             340
Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
                                                 365
                             360
 His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
                                             380
                         375
 Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln
     370
                                         395
                     390
 Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
                                      410
                 405
 Lys Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu
                                                      430
                                  425
             420
 Val Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr
                              440
         435
 <210> 3
 <211> 1737
 <212> DNA
  <213> Nicotiana tabacum
  <220>
  <221> misc_feature
  <222> (733)...(741)
  <223> function: Asn codon in this context is a potential
        glycosylation site;
        product: N-glycosylation consensus sequence;
        phenotype: N-glycans modulate protein properties;
        standard_name: N-glycosylation site;
        label: pot-CHO;
        note: GnTI sequences from animals do not contain
        this feature.
```

Page 5

<221> CDS

<222> (127)...(1467)

```
<223> codon start: 127;
      function: initiates complex N-glycans on secretory
      glycoproteins;
      EC number: 2.4.1.101;
      product: beta-1,2-N-acetylglucosaminlytransferase I;
      evidence: EXPERIMENTAL;
      gene: cgl;
      standard name: gntI;
      label: ORF;
      note: first gntI sequence from tobacco
      (unpublished).
<221> 5'UTR
<222> (15)...(126)
<221> 3'UTR
<222> (1468)...(1723)
<221> CDS
<222> (154)...(213)
<223> function: membrane anchor (amino acids 10-29);
       product: hydrophobic amino acid stretch in GnTI;
       standard_name: membrane anchor of a type II golgi
       protein.
 <221> misc_feature
 <222> (1)...(14)
 <223> function: use for cloning the cDNA library in
       Lambda ZAPII;
       product: EcoRI/NotI-cDNA adapter;
       number: 1.
 <221> misc feature
 \langle 222 \rangle (172\overline{4})...(1737)
 <223> product: EcoRI/NotI-cDNA adapter;
       number: 2.
 gaattcgcgg ccgccattga cttgatccta actgaacagg caaagtaaat ccagcgatga 60
 aacactcata actgaacact gagagactat tcgctttctc ctaaagcctt caatcgaatt 120
 cgcacg atg aga ggg aac aag ttt tgc tgt gat ttc cgg tac ctc ctc
         Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu
  atc ttg gct gct gcc ttc atc tac aca cag atg cgg ctt ttt gcg
                                                                      216
  Ile Leu Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala
                        20
   15
  aca cag tca gaa tat gca gat cgc ctt gct gct gca att gaa gca gaa
                                                                      264
  Thr Gln Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ile Glu Ala Glu
                    35
  aat cat tgt aca agc cag acc aga ttg ctt att gac cag att agc ctg
                                                                      312
  Asn His Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu
                50
  cag caa gga aga ata gtt gct ctt gaa gaa caa atg aag cgt cag gac
                                                                      360
```

Gln Gl		ly <i>1</i> 65	Arg	Ile	Val	Ala	Leu 70	Glu	Gl	u G	Sln	Met	Lys 75	s Ai	rg (Sln	As	р	
cag ga Gln Gl			cga Arg	caa Gln	tta Leu	agg Arg 85	gct Ala	ctt Leu	gt Va	it o	cag Gln	gat Asp 90	cti	z ga u Gi	aa a lu S	agt Ser	aa Ly	g s	408
ggc at Gly II 95	ta a le I	ıaa .ys	aag Lys	ttg Leu	atc Ile 100	gga Gly	aat Asn	gta Val	G.	L11 1	atg Met 105	cca Pro	gt Va	g g l A	ct (gct Ala	gt Va 11	_	456
gtt g Val V	tt a al N	atg Met	gct Ala	tgc Cys 115	aat Asn	cgg Arg	gct Ala	gat Asp	, т	ac yr 20	ctg Leu	gaa Glu	aa Ly	g a s T		att Ile 125	aa Ly	ia /s	504
tcc a Ser I	tc t le 1	tta Leu	aaa Lys 130	tac Tyr	caa Gln	ata Ile	tct Ser	gtt Val	. A	cg la	tca Ser	aaa Lys	ta Ty		ct ro .40	ctt Leu	t† Pl	ic ne	552
ata t Ile S	er (cag Gln 145	gat Asp	gga Gly	tca Ser	cat His	cct Pro) AS	t g p V	tc al	agg Arg	aaq Lys	g ct s Le 15	, u .	gct Ala	ttg Leu	a S	gc er	600
tat q Tyr <i>F</i>	gat Asp 160	cag Gln	ctg Leu	acg Thr	tat Tyr	ato Met	GTI	g ca n Hi	c t s I	tg Leu	gat Asp	tt. Pho		aa d Lu 1	cct Pro	gtg Val	C H	at is	648
act (Thr (gaa Glu	aga Arg	cca Pro	ggg Gl	g gaç 7 Glu 180	те п	g at	t gc e Al	a t a T	tac Tyr	tac Ty: 185	1	a a s I	tt (gca Ala	cgt Arg	•	at is 90	696
tac Tyr	aag Lys	tgg Trp	gca Ala	tto Lei 19	ı Ası	ca Gl:	g ct n Le	g tt u Ph	le	tac Tyr 200	цy.	g ca s Hi	t a s A	at sn	ttt Phe	ago Sei 205		gt Arg	744
gtt Val	atc Ile	ata Ile	cta Lev 210	ı GL	a ga u As	t ga p As	t at p Me	g ga t G1 21	Lu	att Ile	gc Al	c cc a Pr	t g	at	ttt Phe 220		t q e <i>l</i>	gac Asp	792
ttt Phe	ttt Phe	gaç Glu 225	ı Al	t gg a Gl	a gc y Al	a un	t ct r Le 23	eu lit	zu	MO F	<i>, </i>	9 235		-1-	tcç Ser	at Il	t e	atg Met	840
gct Ala	att Ile 240	tc! Se:		t tg r Tr	g aa p As	t ga n As	sp As	at g sn G	ga ly	caa Glr	a at n Me	, c	ag t ln 1 50	tt Phe	gto Val	c ca L Gl	a n	gat Asp	888
cct Pro 255			t ct a Le	t ta u Ty	ac co yr Ai 26	:g 5	ca g er A	at t sp P	tt he	tt! Phe	t co e Pi 20	-0 0	gt ly	ctt Leu	gg: Gl:	a tç y Tr	p g	atg Met 270	936
	tca Ser	a aa c Ly	a to	er Ti	et to nr T: 75	gg ga	ac g sp G	aa t lu I	ta Jeu	tc Se 28		ca a ro L	ag ys	tgg Trp	cc Pr	a aa o Ly 28	•	gct Ala	984
tac Tyr	tg:	g ga o As	ic ga		gg c	ta a eu A	ga c rg I	tc a eu I	aaa Lys	ga Gl	ga uA	at c sn H	ac Iis	aga Arg	gg Gl	t co y A:	ga rg	caa Gln	1032

290		295	300	
ttt att cgc cca Phe Ile Arg Pro 305	gaa gtt tgc aga Glu Val Cys Arg 310	IUT TAT Wall THE		1080
tct agt ttg ggg Ser Ser Leu Gly 320	cag ttt ttc aag Gln Phe Phe Lys 325	cag tat ctt gag Gln Tyr Leu Glu 330	cca att aaa cta Pro Ile Lys Leu	1128
aat gat gtc cag Asn Asp Val Gln 335	gtt gat tgg aag Val Asp Trp Lys 340	tca atg gac ctt Ser Met Asp Leu 345		1176
gag gac aat tac Glu Asp Asn Tyr	gtg aaa cac ttt Val Lys His Phe 355	ggt gac ttg gtt Gly Asp Leu Val 360	aaa aag gct aag Lys Lys Ala Lys 365	1224
ccc atc cat gga Pro Ile His Gly 370	gct gat gct gtc Ala Asp Ala Val	ttg aaa gca ttt Leu Lys Ala Phe 375	aac ata gat ggt Asn Ile Asp Gly 380	1272
gat gtg cgt att Asp Val Arg Ile 385	cag tac aga gat Gln Tyr Arg Asp 390	caa cta gac ttt Gln Leu Asp Phe	gaa aat atc gca Glu Asn Ile Ala 395	1320
cgg caa ttt ggc Arg Gln Phe Gly 400	att ttt gaa gaa Ile Phe Glu Glu 405	a tgg aag gat ggt n Trp Lys Asp Gly 410	· w =	1368
gca tat aaa gga Ala Tyr Lys Gly 415	ata gta gtt tto Ile Val Val Pho 420	c cgg tac caa acg e Arg Tyr Gln Thr 425	g too aga ogt gta Ser Arg Arg Val 430	1416
ttc ctt gtt ggc Phe Leu Val Gly	c cat gat tcg ct 7 His Asp Ser Le 435	t caa caa ctc gga u Gln Gln Leu Gly 440	a att gaa gat act y Ile Glu Asp Thr 445	1464
taa caaagatatg *	attgcaggag cccg	ggcaaa atttttgac†	t tattgggtag	1517
tacggaggag ctc		tc agaacataaa ga	tacaacgt tttaatgtta tcttagta ttggatgaa ggaaatgt agccctgtaa	
<210> 4 <211> 446 <212> PRT <213> Nicotian	a tabacum			
1 Ala Ala Val Al			yr Leu Leu Ile Leu 15 eu Phe Ala Thr Gln 30	

```
Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu Asn His
                            40
Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu Gln Gln
                        55
Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln Asp Gln Glu
                    70
Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile
                                   90
                85
Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val Val
                               105
           100
Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys Ser Ile
                           120
Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser
                       135
Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Asp
                                      155
                   150
Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val His Thr Glu
                                                       175
                                  170
               165
Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys
                               185
           180
Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg Val Ile
                                     205
                            200
 Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Phe Phe
                                          220
                        215
 Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile
                                      235
                    230
 Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp Pro Tyr
                                   250
                245
 Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser
                                                  270
                                265
            260
 Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp
                            280
 Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile
                                           300
                        295
 Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser
                                       315
                     310
 Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp
                                    330
                 325
 Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp
                                                    350
                                345
 Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile
                            360
 His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val
                                            380
                         375
 Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala Arg Gln
                                       395
                     390
  Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr
                                    410 415
                 405
  Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val Phe Leu
                                425
  Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr
                             440
  <210> 5
  <211> 1854
```

<212> DNA

<213> Arabidopsis thaliana

```
<220>
<221> misc feature
<222> (1185)...(1193)
<223> function: Asn Codon is a potential glycosylation
      product: Consensus sequence for N-glycosylation;
      phenotype: N glycans modulate protein
      characteristics;
      standard name: N glycosylation site;
      label: pot-CHO;
      note: absent in animal GnTI sequences.
<221> CDS
<222> (135)...(1469)
<223> codon start: 135;
      function: initiates complex N glycans on secretory
       glycoproteins;
       EC_number: 2.4.1.101;
      product: beta-1,2-N-acetyl glucosaminyl
       transferase I;
       evidence: EXPERIMENTAL;
       gene: cgl;
       standard name: gntI;
       label: ORF;
       note: first gntI sequence from Arabidopsis
       (unpublished).
 <221> 5'UTR
 <222> (19)...(134)
 <221> 3'UTR
 <222> (1470)...(1848)
 <221> CDS
 <222> (157)...(215)
 <223> function: membrane anchor (amino acids 8-27);
       product: hydrophobic amino-acid region in GnTI;
       standard_name: membrane anchor of a Type II Golgi
        protein;
        note: identified by comparison with animal GnTI
        sequences.
  <221> misc feature
  <222> (1)...(18)
  <223> function: for preparation of a cDNA library in
        Lambda ACT;
        product: XhoI-cDNA-Adaptor;
        number: 1.
  <221> misc feature
  \langle 222 \rangle (184\overline{9}) \dots (1854)
  <223> product: XhoI-cDNA-Adaptor;
        number: 2.
  ctcgaggcca cgaaggccac cgtttttgtt ataacgaacg acaccgtttc aaacaacttc 60
  cttattaget ageteettee eggeggeaaa caccagaaga tecacegett ttgatetggt 120
```

tgtttgtcgt cgat	atg gcg Met Ala 1	agg atc Arg Ile	tcg Ser 5	tgt (Cys <i>l</i>	gac t Asp I	tg a Leu i	aga t Arg l	ttt Phe 10	ctt Leu	ctc Leu	170
atc ccg gca gct Ile Pro Ala Ala 15	ttc atg Phe Met	ttc atc Phe Ile 20	tac Tyr	atc (Ile (cag a Gln N	atg Met	agg (Arg :	ctt Leu	ttc Phe	cag Gln	218
acg caa tca cag Thr Gln Ser Gln 30	tat gca Tyr Ala	gat cgc Asp Arg 35	ctc Leu	agt Ser	tcc (Ser <i>i</i>	gct Ala 40	atc Ile	gaa Glu	tct Ser	gag Glu	266
aac cat tgc act Asn His Cys Thr 45	agt caa Ser Gln 50	atg cga Met Arg	ggc Gly	ctc Leu	ata Ile 55	gat Asp	gaa Glu	gtt Val	agc Ser	atc Ile 60	314
aaa cag tcg cgg Lys Gln Ser Arg	att gtt Ile Val 65	gcc ctc Ala Leu	gaa Glu	gat Asp 70	atg Met	aag Lys	aac Asn	cgc Arg	cag Gln 75	gac Asp	362
gaa gaa ctt gtg Glu Glu Leu Val 80	Gln Leu	aag gat Lys Asp	cta Leu 85	atc Ile	cag Gln	acg Thr	ttt Phe	gaa Glu 90	aaa Lys	aaa Lys	410
gga ata gca aaa Gly Ile Ala Lys 95	a ctc act s Leu Thr	caa ggt Gln Gly 100	y Gly	cag Gln	atg Met	cct Pro	gtg Val 105	gct Ala	gct Ala	gta Val	458
gtg gtt atg gc Val Val Met Ala 110	c tgc agt a Cys Ser	cgt gca Arg Ala	a gac a Asp	tat Tyr	ctt Leu	gaa Glu 120	agg Arg	act Thr	gtt Val	aaa Lys	506
tca gtt tta ac Ser Val Leu Th 125	a tat caa r Tyr Glr 130	Thr Pr	c gtt o Val	gct Ala	tca Ser 135	aaa Lys	tat Tyr	cct Pro	cta Leu	ttt Phe 140	554
ata tct cag ga Ile Ser Gln As	t gga tct p Gly Ser 145	gat ca Asp Gl	a gct n Ala	gtc Val 150	гаг	agc Ser	aag Lys	tca Ser	tto Leu 155	1 501	602
tat aat caa tt Tyr Asn Gln Le 16	u Thr Ty	atg ca Met Gl	g cac n His 165	Leu	gat Asp	ttt Phe	gaa Glu	cca Pro 170	va.	g gtc L Val	650
act gaa agg co Thr Glu Arg Pr 175	t ggt gaa o Gly Gl	a ctg ac u Leu Th 18	ir Ala	tac Tyr	tac Tyr	aag Lys	att Ile 185	HIC	cgi Are	t cac g His	698
tac aag tgg go Tyr Lys Trp Al 190	a ctg ga .a Leu As	c cag tt p Gln Le 195	ig ttt eu Phe	tac Tyr	aaa Lys	cac His 200	ь пля	ttt Phe	ag e Se	t cga r Arg	746
gtg att ata ct Val Ile Ile Le 205	a gaa ga eu Glu As 21	p Asp Me	ig gaa et Gli	a att ı Ile	gct Ala 215	z ET	a gad o Asp	c tto Pho	c tt e Ph	t gat e Asp 220	794
tac ttt gag g	ct gca gc	t agt c	tc at	g gat	t ago	g ga	t aaa	a ac	c at	t atg	842

Tyr	Phe	Glu	Ala	Ala 225	Ala	Ser	Leu	Met	Asp 230	Arg	Asp	Lys	Thr	Ile 235	Met	
gct Ala	gct Ala	tca Ser	tca Ser 240	tgg Trp	aat Asn	gat Asp	aat Asn	gga Gly 245	cag Gln	aag Lys	cag Gln	ttt Phe	gtg Val 250	cat His	gat Asp	890
ccc Pro	tat Tyr	gcg Ala 255	cta Leu	tac Tyr	cga Arg	tca Ser	gat Asp 260	ttt Phe	ttt Phe	cct Pro	ggc Gly	ctt Leu 265	gly ggg	tgg Trp	atg Met	938
ctc Leu	aag Lys 270	aga Arg	tcg Ser	act Thr	tgg Trp	gat Asp 275	gag Glu	tta Leu	tca Ser	cca Pro	aag Lys 280	tgg Trp	cca Pro	aag Lys	gct Ala	986
tac Tyr 285	tgg Trp	gat Asp	gat Asp	tgg Trp	ctg Leu 290	aga Arg	cta Leu	aag Lys	gaa Glu	aac Asn 295	cat His	aaa Lys	ggc Gly	cgc Arg	caa Gln 300	1034
ttc Phe	att Ile	gca Ala	ccg Pro	gaa Glu 305	gtc Val	tgt Cys	aga Arg	aca Thr	tac Tyr 310	aat Asn	ttt Phe	ggt Gly	gaa Glu	cat His 315	ggg Gly	1082
tct Ser	agt Ser	ttg Leu	gga Gly 320	cag Gln	ttt Phe	ttc Phe	agt Ser	cag Gln 325	tat Tyr	ctg Leu	gaa Glu	cct Pro	ata Ile 330	aag Lys	cta Leu	1130
aac Asn	gat Asp	gtg Val 335	Thr	gtt Val	gac Asp	tgg Trp	aaa Lys 340	gca Ala	aag Lys	gac Asp	ctg Leu	gga Gly 345	Tyr	ctg Leu	aca Thr	1178
gag Glu	gga Gly 350	Asn	tat Tyr	acc Thr	aag Lys	tac Tyr 355	Phe	tct Ser	ggc	tta Leu	gtg Val 360	Arg	caa Gln	gca Ala	cga Arg	1226
cca Pro 365	Ile	caa Gln	ggt Gly	tct Ser	gac Asp 370	Leu	gtc Val	tta Leu	aag Lys	gct Ala 375	GIn	aac Asn	ata Ile	aag Lys	gat Asp 380	1274
gat Asp	gat Asp	cgt Arç	ı Ile	e Arc	, Tyr	Lys	gac Asp	Gln	. Val	. GIV	j ttt i Phe	gaa Glu	cgc Arg	att Ile 395	gca Ala	1322
Gl ⁷	g gaa 7 Glu	a ttt ı Phe	ggt Gly 400	/ Ile	a ttt e Phe	gaa Glu	a gaa ı Glu	tgg Trp 405	ь гла	g gat s Asp	ggt Gly	gto Y Val	g cca L Pro 410	HIG	a aca g Thr	1370
gca Ala	a tat a Tyi	t aaa t Lys 41!	s Gl	a gta y Val	a gto L Val	g gto Va	g ttt l Phe 420	e Arg	a ato	c caq e Gli	g aca	a acar Th:	r Arc	a cgt g Arg	g ta	1418
tto Phe	c cto e Leo 430	ı Va	t gg l Gl	g cca y Pro	a gat o Asp	t to Se:	r Va.	a atq l Mei	g ca CGl	g cti n Lei	t gga u Gl	À ТТ	t cga e Arq	a aat g Asi	t tcc n Ser	1466
tga			acat	atg	aaag	gaa	aaga	agat [.]	tt t	ggac	cgca	t gc	agcc	tcct		1519

```
tctagcagct gttaggttgt attgttattt atggatgagt ttgtagagcg gtggggttaa 1579
ctttaacagc aaggaagete tggtgaccag getgattgge ttagaagtta tgggaaccec 1639
ttgaaagggt cagggttaaa tatatttcag ttgttttatt agtgattatc ttgtgggtaa 1699
cttatacgaa tgcaaatcat tctatgcagt ttttcttcgt cccacttgtt ttggcttctc 1759
tattgctagt gtacatatct cttcaaacat gtactaaata atgcgtgttg cttcaaagaa 1819
gtaactttta ttaaaaaaaa aaaaaaaaac tcgag
<210> 6
<211> 444
<212> PRT
<400> 6
Met Ala Arg Ile Ser Cys Asp Leu Arg Phe Leu Leu Ile Pro Ala Ala
                                      10
Phe Met Phe Ile Tyr Ile Gln Met Arg Leu Phe Gln Thr Gln Ser Gln
                                                      30
                                  25
             20
Tyr Ala Asp Arg Leu Ser Ser Ala Ile Glu Ser Glu Asn His Cys Thr
                              40
         35
Ser Gln Met Arg Gly Leu Ile Asp Glu Val Ser Ile Lys Gln Ser Arg
     50
Ile Val Ala Leu Glu Asp Met Lys Asn Arg Gln Asp Glu Glu Leu Val
                                          75
                      70
Gln Leu Lys Asp Leu Ile Gln Thr Phe Glu Lys Lys Gly Ile Ala Lys
                                                           95
                                      90
                  85
Leu Thr Gln Gly Gln Met Pro Val Ala Ala Val Val Met Ala
                                 105
             100
Cys Ser Arg Ala Asp Tyr Leu Glu Arg Thr Val Lys Ser Val Leu Thr
                             120
         115
Tyr Gln Thr Pro Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln Asp
                                             140
                         135
 Gly Ser Asp Gln Ala Val Lys Ser Lys Ser Leu Ser Tyr Asn Gln Leu
                                         155
                     150
 Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val Val Thr Glu Arg Pro
                                                          175
                                      170
                 165
 Gly Glu Leu Thr Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp Ala
                                                      190
                                 185
             180
 Leu Asp Gln Leu Phe Tyr Lys His Lys Phe Ser Arg Val Ile Ile Leu
                              200
         195
 Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Tyr Phe Glu Ala
                                              220
                          215
 Ala Ala Ser Leu Met Asp Arg Asp Lys Thr Ile Met Ala Ala Ser Ser
                                          235
                      230
 Trp Asn Asp Asn Gly Gln Lys Gln Phe Val His Asp Pro Tyr Ala Leu
                                                          255
                                      250
                  245
 Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Lys Arg Ser
                                                      270
                                  265
             260
 Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp Asp
                                                  285
                              280
         275
 Trp Leu Arg Leu Lys Glu Asn His Lys Gly Arg Gln Phe Ile Ala Pro
                                              300
                          295
 Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu Gly
                                                               320
                                          315
                      310
 Gln Phe Phe Ser Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val Thr
                                                           335
                                      330
 Val Asp Trp Lys Ala Lys Asp Leu Gly Tyr Leu Thr Glu Gly Asn Tyr
                                                       350
                                  345
              340
```

١

```
Thr Lys Tyr Phe Ser Gly Leu Val Arg Gln Ala Arg Pro Ile Gln Gly
                             360
       355
Ser Asp Leu Val Leu Lys Ala Gln Asn Ile Lys Asp Asp Asp Arg Ile
                         375
Arg Tyr Lys Asp Gln Val Glu Phe Glu Arg Ile Ala Gly Glu Phe Gly
                                         395
                    390
Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Thr Ala Tyr Lys Gly
                                                         415
                                     410
                405
Val Val Phe Arg Ile Gln Thr Thr Arg Arg Val Phe Leu Val Gly
                                 425
            420
Pro Asp Ser Val Met Gln Leu Gly Ile Arg Asn Ser
                             440
<210> 7
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer for cloning plant genes
<221> misc feature
 <222> (6)...(12)
 <223> n at positions 6, 9 & 12 = inosine.
 <400> 7
                                                                     24
 tgygynwsng cntggmayga yaay
 <210> 8
 <211> 24
 <212> DNA
 <213> Artificial Sequence
 <223> Primer for cloning plant genes
 <221> misc feature
 <222> (4)...(10)
 <223> n at positions 4, 7, & 10 = inosine.
 <221> misc feature
 <222> 13
  \langle 223 \rangle n = a, c, g, or t.
  <400> 8
                                                                      24
 ccancentrn cengsraara arte
  <210> 9
  <211> 14
  <212> PRT
  <213> Artificial Sequence
  <220>
  <223> Primer for cloning plant genes
  <221> VARIANT
  <222> 2
```

```
\langle 223 \rangle Xaa = Arg or Met.
<221> VARIANT
<222> 9
\langle 223 \rangle Xaa = Asp or Tyr.
<400> 9
Gln Xaa Gln Phe Val Gln Asp Pro Xaa Ala Leu Tyr Arg Ser
                  5
<210> 10
<211> 48
<212> DNA
<213> Artificial Sequence
<223> Primer for cloning plant genes
atcggaaagc ttggatcccc agtggcrgct gtagttgtta tggcttgc
                                                                       48
<210> 11
<211> 27
<212> DNA
<213> Artificial Sequence
<223> Primer for cloning plant genes
<400> 11
                                                                       27
ggcccccct cgaggtcgac ggtatcg
 <210> 12
 <211> 51
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Primer for cloning plant genes
 gggcctctag actccagage yactactctt ccttgctgct ggctaatctt g
 <210> 13
 <211> 40
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Primer for cloning plant genes
 <400> 13
                                                                        40
 catggatccc tcgagaagcg tcaggaccag gagtgccggc
 <210> 14
  <211> 37
```

<212> DNA <213> Artificial Sequence <220> <223> Primer for cloning plant genes <400> 14 atcccgggat ccgctacgta tcttcaactc caagttg

37